



**University of  
Zurich**<sup>UZH</sup>

**Zurich Open Repository and  
Archive**

University of Zurich  
University Library  
Strickhofstrasse 39  
CH-8057 Zurich  
[www.zora.uzh.ch](http://www.zora.uzh.ch)

---

Year: 2015

---

## **Factors influencing the detectability of early warning signals of population collapse**

Clements, Christopher F ; Drake, John M ; Griffiths, Jason I ; Ozgul, Arpat

**Abstract:** The recent description of potentially generic early warning signals is a promising development that may help conservationists to anticipate a population's collapse prior to its occurrence. So far, the majority of such warning signals documented have been in highly controlled laboratory systems or theoretical models. Data from wild populations, however, are typically restricted both temporally and spatially due to limited monitoring resources and intrinsic ecological heterogeneity – limitations that may affect the detectability of generic early warning signals, as they add additional stochasticity to population abundance estimates. Consequently, spatial and temporal subsampling may serve either to muffle or magnify early warning signals. Using a combination of theoretical models and analysis of experimental data, we evaluate the extent to which statistical warning signs are robust to data corruption.

DOI: <https://doi.org/10.1086/681573>

Posted at the Zurich Open Repository and Archive, University of Zurich

ZORA URL: <https://doi.org/10.5167/uzh-118908>

Journal Article

Published Version

Originally published at:

Clements, Christopher F; Drake, John M; Griffiths, Jason I; Ozgul, Arpat (2015). Factors influencing the detectability of early warning signals of population collapse. *The American Naturalist*, 186(1):50-58.

DOI: <https://doi.org/10.1086/681573>

# Factors Influencing the Detectability of Early Warning Signals of Population Collapse

Christopher F. Clements,<sup>1,\*</sup> John M. Drake,<sup>2</sup> Jason I. Griffiths,<sup>3</sup> and Arpat Ozgul<sup>1</sup>

1. Institute of Evolutionary Biology and Environmental Studies, University of Zurich, Zurich CH-8057, Switzerland; 2. Odum School of Ecology, University of Georgia, Athens, Georgia 30602; 3. Department of Animal and Plant Sciences, University of Sheffield, Western Bank, Sheffield S10 2TN, United Kingdom

Submitted August 22, 2014; Accepted January 22, 2015; Electronically published May 7, 2015

Online enhancements: appendixes, zip file.

**ABSTRACT:** The recent description of potentially generic early warning signals is a promising development that may help conservationists to anticipate a population's collapse prior to its occurrence. So far, the majority of such warning signals documented have been in highly controlled laboratory systems or in theoretical models. Data from wild populations, however, are typically restricted both temporally and spatially due to limited monitoring resources and intrinsic ecological heterogeneity—limitations that may affect the detectability of generic early warning signals, as they add additional stochasticity to population abundance estimates. Consequently, spatial and temporal subsampling may serve to either muffle or magnify early warning signals. Using a combination of theoretical models and analysis of experimental data, we evaluate the extent to which statistical warning signals are robust to data corruption.

**Keywords:** conservation, critical slowing down, observation error, search effort, subsampling, wild populations.

## Introduction

Populations approaching certain kinds of thresholds have been shown to exhibit generic early warning signals—leading indicators—prior to a transition, a function of a population's diminishing ability to return to its prior state after environmental perturbation, known as critical slowing down (Contamin and Ellison 2009; Drake and Griffen 2010a; Carpenter et al. 2011; Dakos et al. 2011, 2012b). Such early warning signals include an increase in the standard deviation of a population's size, a decrease in the return rate of a population to its mean after a perturbation, and an increase in the autocorrelation of a population's demographic fluctuations (Dakos et al. 2012a). These generic leading indicators arise from mathematical properties of stochastic dynamical systems as they approach a tipping point and appear to be present in vastly different systems, including stock

markets, global climatic change, and populations (for a review, see Scheffer et al. 2009). The development and testing of these statistical early warning signals suggests that a predictive framework for population risk might be developed that requires only a relatively limited amount of demographic data (Hefley et al. 2013; Krkošek and Drake 2014). From a conservation perspective, such a framework could provide an important tool for prioritizing conservation efforts and funding to target those species approaching a critical transition. However, while there is a large body of work dedicated to collecting abundance data from wild populations (see, e.g., Collen et al. 2009; NERC Centre for Population Biology 2010), it remains unclear whether critical slowing down may be detected in such data. If not, future work should concentrate on developing metrics of population risk that are tailored to the data available.

Tests of the detectability of statistical early warning signals have typically used simulated data or laboratory experiments. Drake and Griffen (2010a) used a long-term experiment (approximately 60 generations; Ebert 2005) in small-scale aquatic microcosms to show that leading indicators of population collapse could be identified several generations before populations became extinct. Field tests of such signals have been provided by, among others, Carpenter et al. (2011), whose whole-lake manipulation demonstrated that statistical signals of a regime shift were detectable more than a year before the shift occurred. Both of these experimental manipulations highlight the challenges faced when collecting data for detecting critical slowing down; in Drake and Griffen (2010a), populations were sampled in triplicate once per week (approximately once per generation; Ebert 2005) for over a year, while Carpenter et al. (2011) collected data on three trophic levels daily during the summer sampling period for 3 years. These and other studies (e.g., Dakos et al. 2008) suggest that detecting early warning signals may require high-resolution, high-quality data—potentially unavailable in conservation practice. Because early warning signals are derived from statis-

\* Corresponding author; e-mail: chriscllementsresearch@gmail.com.

Am. Nat. 2015. Vol. 186, pp. 50–58. © 2015 by The University of Chicago. 0003-0147/2015/18601-55709\$15.00. All rights reserved.  
DOI: 10.1086/681573

tical signals of time series, the quality of the data available is critical to the utility of these metrics of population risk. Consequently, factors that affect signal-to-noise ratio could significantly affect the usefulness of leading indicators. Indeed, previous work has shown that, for example, the age structure of a population may obscure statistical early warning signals (Krkošek and Drake 2014).

Data on the demography of an increasing number of wild populations are becoming available for terrestrial, marine, and freshwater environments (Loh et al. 2005; Collen et al. 2009; NERC Centre for Population Biology 2010). Indeed, the analysis of global data sets of the change in population abundance through time is one of the methods used to analyze the current state of global biodiversity (e.g., the Living Planet Index; Collen et al. 2009). However, such data sets highlight some of the limitations of the available data: data are often limited both spatially and temporally because of the inherent problems of sampling wild populations.

Spatial limitations on the sampling of wild populations result from the impossibility of sampling 100% of a species' habitat. Typically, a subset of the population is counted (say, along a transect or within a defined area), with this subset being assumed to reflect the true (total) population abundance (e.g., Franzetti et al. 2012; but see Hefley et al. 2013). However, in itself, this subsampling generates uncertainty, as chance events such as the movement of organisms, the heterogeneous distribution of individuals, and the skill of the observers influence recorded population abundance (Muhlfeld et al. 2006). While there are a variety of sophisticated methods that seek to address the problems raised by spatial subsampling (e.g., Wright et al. 2009; Ford et al. 2012), the uncertainty generated at the sampling stage can never be fully resolved.

Similarly, temporal limitations on data are imposed by infrequent sampling of a population's abundance through time caused by, for example, limited funding or limited available personnel (ter Braak et al. 1994; Clark and Bjørnstad 2004; Birkhead 2014). This may take the form of time series that cover a short period with relatively frequent sampling, time series that cover a longer period of time with more infrequent sampling, or irregular sampling through time. Irregular temporal sampling is of particular concern, as many statistical leading indicators require data that are sampled at regular intervals (Dakos et al. 2012a). Irregularly sampled data must therefore be interpolated to create evenly spaced data, which may introduce artificial autocorrelation (Boettiger and Hastings 2012).

Given the commonality of these limitations, it would be useful to determine how the subsampling that is employed while monitoring wild populations affects both the prevalence and the strength of early warning signals. Previous work has suggested that generic statistical indicators will

perform poorly when observational errors are large (Dai et al. 2012; Ives and Dakos 2012; but see Hefley et al. 2013). One method that may be used to ameliorate some of these issues would be the use of state-space models (which allow the data collection process to be incorporated into dynamic population models of a species, where the model can pass through a critical transition), which could explicitly account for observational errors (Hefley et al. 2013). Previous work has used this approach to detect potential transcritical bifurcations in wild populations (Hefley et al. 2013). However, this approach is significantly more complicated to implement and more problem specific than the methods of, say, Dakos et al. (2008) and Scheffer et al. (2009). Moreover, Hefley et al. (2013) found that even with significant observational error generated by spatial subsampling, statistical leading indicators may still perform reasonably well. Given this additional complexity and the apparent robustness of statistical approaches, state-space model approaches are not considered further here. However, identifying whether generic statistical leading indicators are suitable for use with spatially and temporally limited wild population data is a key issue.

We used a combination of simulation modeling and analysis of experimental data to investigate how the percentage of the habitat searched to assess population abundances and the temporal frequency of sampling alter the presence and strength of generic early warning signals. To address this, we first simulated data using a model previously shown to exhibit early warning signals prior to a regime shift (Dakos et al. 2012a) and then corrupted this perfect data by simulating search regimes and search frequencies. We identified which early warning signals are most robust to subsampling of the data and what portion of the habitat must be searched (or frequency of sampling must be employed) before trends in early warning signals were reliably identified. We then duplicated this analysis using data from an experiment where early warning signals had previously been identified prior to population collapse (Drake and Griffen 2010b). Finally, based on these results, we made recommendations as to which early warning signals should be considered applicable for data collected on wild population abundances.

## Methods

### *Simulated Time Series*

Following Dakos et al. (2012a), we simulated a time series in which a critical transition is known to have occurred, driven by a slow forcing, using the stochastic differential equation

$$dx = \left[ rx \left( 1 - \frac{x}{K} \right) - c \frac{x^2}{x^2 + h^2} \right] dt + \sigma x dW$$

that uses the Euler approximation with an Euler time step size ( $\delta t$ ) of 0.025 (see zip file, available online, where code to implement this is provided). The population's growth rate is  $r$ ,  $K$  is the carrying capacity,  $h$  is the half-saturation constant,  $c$  is the harvesting rate, and  $dW$  is a normally distributed (white Gaussian noise process) stochastic perturbation that was applied at each time step with expectation  $\mu(x_i, t)$  and noise intensity  $\sigma(x_i, t)^2$ . Thus, at each Euler time step, population rates of change were calculated based on the above model and were used to forecast the expected abundance at time  $t + \delta t$ . Process variation was then added to the prediction by allowing the realized  $Nt + \delta t$  to be sampled from a normal distribution with a mean equal to the expected value and a standard deviation of  $\sqrt{\sigma}$ .

This model produces a similar, although not identical, regime shift to the experimental data described below—the model exhibits a fold catastrophe (where a system transitions discontinuously from one state to another), whereas the experimental data exhibit a transcritical bifurcation (where a population changes continuously to extinction at a cusp). Both critical transitions can generate leading indicators of collapse (Drake and Griffen 2010a; Dakos et al. 2012a).

Unlike Dakos et al. (2012a), we simulated the population abundance at each time step (i.e., the number of individuals of a population rounded to the nearest whole number, rather than biomass, which changes on a continuous scale) to enable resampling to reflect a real observation process and more naturally represent the experimental data of Drake and Griffen (2010a; see below). In this model,  $c$  can be regarded as a rate of harvesting; for example, the removal of individuals while fishing. As in Dakos et al. (2012a), we simulated data for 1,000 time steps. The speed at which a population approaches a critical transition is influenced by the rate of change of the system from a steady state. In our case, this is largely governed by the parameter  $c$ . As this approach speed could alter the prevalence of early warning signals (Brock and Carpenter 2010; Krkošek and Drake 2014), we simulated 10 time series where the value of  $c$  increased linearly (from 1) at different rates (0.025, 0.03, 0.035, ..., 0.07  $c$  per time step; fig. 1a). Five replicates were simulated for each rate of change of  $c$  (fig. 1a).

### Experimental Data

In addition to the simulated data, we analyzed an experimental data set where signals of critical slowing down have previously been shown to occur (Drake and Griffen 2010a). Drake and Griffen (2010a) used replicate populations of the freshwater cladoceran *Daphnia magna* to generate long-term (413 days, ~60 generations; Ebert 2005) population abundance data, sampling populations for abun-

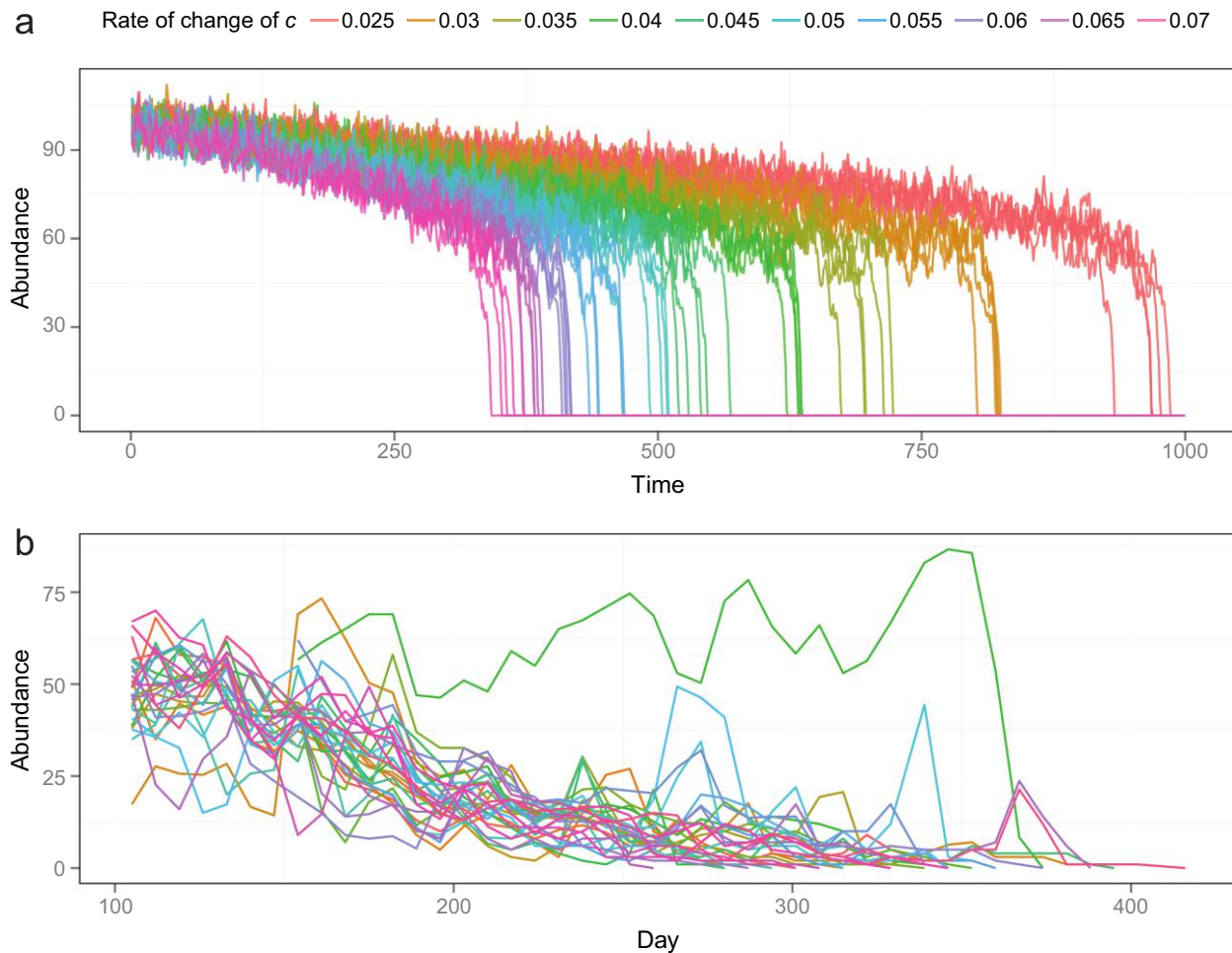
dance once per week. The original study had one treatment group, in which environmental conditions deteriorated (declining resource input), causing extinction, and one control group, in which environmental conditions were constant. Our study used data from the deteriorating environment treatment only (where early warning signals would be expected to occur). Each population's time series was treated independently, with a total of 30 replicates where habitat quality had been degraded through time. All 30 populations fell to extinction during the experimental period, with the fastest collapse occurring by day 259 and the slowest by day 416. As in Drake and Griffen (2010a), we excluded the transient dynamics between days 0 and 105 and analyzed the variation in population abundances from day 100 to the sampling date immediately preceding extinction (1 week, or approximately one generation [Ebert 2005], prior to extinction).

### Resampling the Data

Based on Clements et al. (2014), the simulated and experimental data were then resampled as follows: to simulate the effects of imperfect spatial sampling of a population, a process of constant search effort was applied to each record of population abundance through time. Search effort (the proportion of a habitat searched) was simulated from 0.01 to 1 in increments of 0.01 for the simulated data. Because of the generally lower abundances of the experimental data, search efforts from 0.1 to 1 in increments of 0.01 were implemented. The number of observed individuals at each time step was then calculated by drawing a binomial random variate with size equal to the total abundance and probability set to the search effort, and thus populations were assumed to be randomly distributed across space. It should be noted that such subsampling represents not only imperfect spatial sampling but also the imperfect detection of individuals—for example, searching 100% of a habitat and observing 10% of the individuals (a detection probability of 0.1) is approximately equivalent to searching 10% of a habitat but observing 100% of the individuals within that area (a detection probability of 1), assuming those individuals are homogeneously or randomly distributed. Thus, our results can be interpreted as the effects of either less than perfect spatial sampling or imperfect detection or a combination of both. Simulations were run 10 times for each level of search effort for each replicate of the model and experimental population.

To simulate infrequent temporal sampling, we manipulated the number of time points in each time series by randomly selecting a subset of each perfect time series (i.e., without any spatial sampling) to be used in the calculation of early warning signals. Proportions of 0.1 to 1 (in 0.01 increments) of the original data set were used, and again





**Figure 1:** *a*, Simulated time series that experiences a critical transition from an underexploited state to an overexploited state. The rate of change of the parameter  $c$  (the harvesting parameter) was determined when this critical transition occurred. *b*, Population data from Drake and Griffen (2010b) that were shown to exhibit early warning signals of population collapse.

simulations were run 10 times for each replicate model and experimental population.

#### *Calculation of Early Warning Signals*

We used the package “earlywarnings” for the statistical software R (R Development Core Team 2014) to calculate the change in five generic early warning signals through time: autocorrelation at the first lag (similarity of value at time  $t$  to the value at  $t - 1$ ), density ratio (ratio of low frequencies to high frequencies within a rolling window), first-order autoregressive coefficient ( $\text{ar}(1)$ ), return rate (rate of return to equilibrium after a perturbation), and standard deviation of the data. Coefficient of variation was not calculated, as the data from the model simulation and experiment were detrended using Gaussian detrending, while kurtosis and skewness were not included, as kurtosis is

expected to be present in data where population flicker between two stable states and skewness provides much less reliable indications of an approaching transition (Dakos et al. 2012a). A rolling window size of 50% of the data sets was used for all the analyses, with a default bandwidth estimated using Silverman’s rule (0.9 times the minimum of the standard deviation and the interquartile range, divided by 1.34 times the sample size to the negative one-fifth power; see earlywarnings help files). In addition, where the temporal frequency of samples was manipulated (and therefore abundance counts were not uniformly distributed through time), the data were linearly interpolated (using the interpolate function in the earlywarnings package). These early warning signals were calculated on a subset of the experimental and simulated data: up until five time points before a population’s collapse for the simulated data and to the sample point prior to extinction for the experimental

data (1 week prior to extinction). Following Dakos et al. (2008), we used Kendall's  $\tau$  to assess trends in leading indicators prior to population collapse. We then analyzed the significance and strength of these correlations to evaluate the relative performance of different potential early warning signals in the data. This approach to forecasting critical transitions has been criticized for being statistically inefficient (Boettiger and Hastings 2012). We note that our goal was not to obtain the most efficient statistical estimator but rather to measure the effect of data corruption on the estimators currently enjoying the widest use. Accordingly, we report our results in terms of trends—for instance, whether or not there is a correlation between a candidate leading indicator and time and whether this trend is in the theoretically expected direction. Thus, for example, if there was a positive correlation between time and standard deviation, this is described as a positive trend, as the standard deviation of a population is expected to increase prior to a bifurcation. Conversely, for return rate—which is expected to decrease as a population approaches a bifurcation—a negative correlation is described as a positive trend in the leading indicator. The strength of these trends was recorded as the Kendall  $\tau$  value. For the experimental data, results presented are for the subset that showed trends in the uncorrupted data sets. For simulated time series, results for the scenario where the parameter  $c$  increased 0.025 per time step are presented in the main text, with results for the other nine simulated data sets reported in appendix B (apps. A and B available online).

## Results

### *Effects of Search Effort*

We show that the correlation between a candidate leading indicator and the time remaining before a critical transition is breached is significantly and substantially affected by the fraction of the habitat searched (fig. 2a). For the five early warning statistics (autocorrelation at first lag, autoregressive coefficient  $\text{ar}(1)$ , density ratio, return rate, and standard deviation), the strength of positive trends in leading indicators increased significantly as the fraction of the habitat searched was increased, with a subsequent decrease in the variation in Kendall  $\tau$  values (fig. 2a). The precise relationship between search effort and strength of trends in leading indicators varied among metrics, with some (autocorrelation at first lag, autoregressive coefficient  $\text{ar}(1)$ , and density ratio) tending to produce a high proportion of strong positive trends, even at low search efforts, whereas others (standard deviation, return rate) exhibited this only when search efforts were greater than  $\sim 50\%$  (fig. 2a). These results appear to hold even when the starting population size of the modeled population is altered by two orders of magnitude

(app. A). The rate of forcing of the system (rate of change of the parameter  $c$ ) had relatively little impact on the relationship between search effort and trends in leading indicators, with all leading indicators showing approximately the same relationship across the different rates of forcing (app. B).

Results for the experimental data were similar. We found strong positive correlations in leading indicators to be more prevalent with higher search efforts. However, search effort appeared to have less of an effect on the strength of these signals compared with simulated data (fig. 2b). Of the five leading indicators tested, return rate produced the least reliable indications of an approaching transition, with only 39% of simulations producing a trend in the direction expected by theory. The other four indicators all showed positive correlations in a large number of simulations ( $>84\%$ ; fig. 2b).

### *Effects of the Frequency of Sampling*

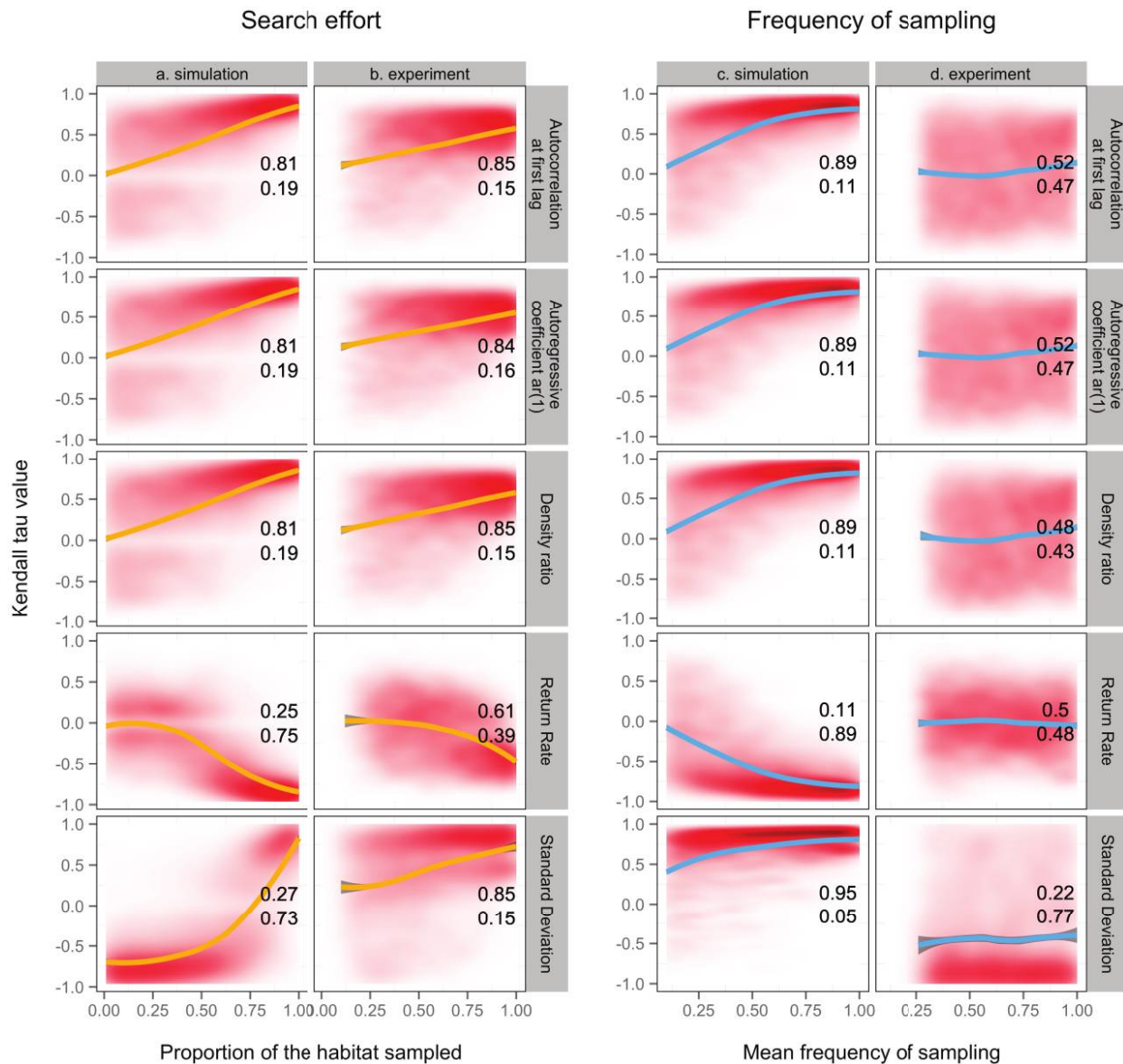
Increasing the mean frequency of sampling increased the strength and number of positive trends in leading indicators (fig. 2c). When there was infrequent sampling (between 0.1 and 0.25 samples per time step), there was considerable variation in the strength of these trends, with a high rate of type II error (failure to detect an approaching critical transition) in all early warning metrics barring standard deviation (fig. 2c). Again, these results appear to hold even when the starting population size of the modeled population is altered dramatically (app. A).

There was weak evidence for an interaction between the rate of change of  $c$  and the mean frequency of sampling (app. B). In general, the strength and direction of leading indicators showed a similar pattern across the different rates of forcing, with small differences between the populations with different rates of change of  $c$ —those with slow rates of change (e.g., 0.03 per time step) tended to have slightly stronger trends in leading indicators than those with fast rates of change of  $c$  (app. B).

Unlike search effort, altering the frequency of sampling of the experimental data did not produce similar results to the model data (fig. 2c, 2d)—the experimental data showed much less frequent positive correlation in leading indicators (trend lines; fig. 2c, 2d), often with large type II error. This was especially prevalent with standard deviation, where 84% of simulations showed a decrease in the standard deviation through time rather than (as expected by theory) an increase (fig. 2d).

## Discussion

The detection of early warning signals prior to population collapse is a potentially powerful tool for prioritizing con-



**Figure 2:** Effects of spatial and temporal subsampling on the strength of early warning signals, measured as the correlation between each leading indicator and time (Kendall  $\tau$  value), using data from a simulated population (with a rate of change of the parameter  $c$  of 0.025; fig. 1) and from Drake and Griffen (2010b). Intensity of the shading shows the distribution of points produced by the simulations in two-dimensional space. Theory suggests that all leading indicators should show positive correlations with time (positive Kendall  $\tau$  values), with the exception of return rate, which should show a negative correlation (negative Kendall  $\tau$  value). Numbers in the upper and lower half of each panel indicate the portion of simulations with a positive or negative trend in Kendall's  $\tau$  correlation coefficient. Lines on each panel show the trends in the values of Kendall  $\tau$  values with search effort and frequency (loess regression); black shaded areas indicate 95% CI.

servation effort and funding. Generic early warning signals may be of particular interest, as they require relatively simple data (population abundance through time) and have been shown to be exhibited, theoretically, by many different kinds of dynamical systems. Here, we show that data limitations imposed by imperfect sampling, an omnipres-

ent factor in wild population surveys, can alter the detectability and reliability of these signals.

The ubiquitous subsampling, both spatial and temporal, associated with the monitoring of wild populations creates unique obstacles to understanding their dynamics (Clark and Bjørnstad 2004; Clements et al. 2013). Such problems

must be overcome, however, as it is impractical to perfectly monitor any population in space and time. Such issues are particularly relevant to the detection of early warning signals of population collapse, as limited spatial sampling and limited temporal sampling can mask the population fluctuations on which generic early warning signals are based.

The results presented here suggest that spatial subsampling significantly alters the detectability of early warning signals and that at low search efforts such early warning signals are liable to be masked by demographic noise. Of the five leading indicators tested here, autocorrelation at first lag, first-order autoregressive coefficient, and density ratio are the most robust to varying search efforts (fig. 2). This supports previous work that has suggested that autocorrelation provides more reliable and robust early warning signals than other metrics (Dakos et al. 2012b; Krkošek and Drake 2014). While it is encouraging that such methods may be robust to what can be considered a large amount of uncertainty, high search efforts to assess a population's abundance may in many instances be unrealistic (Green and Young 1993). Given the prevalence of type II error (erroneously assuming a population is not approaching a critical transition) at low search efforts, explicitly including search effort among leading indicators could prove particularly useful. Previously, something similar had been achieved using state-space models (Hefley et al. 2013), but future work could also seek to include imperfect sampling in the calculation of statistical leading indicators. Alternatively, more powerful model selection approaches (proposed by Boettiger and Hastings [2012]) may be employed, which could prove less susceptible to degraded data. Such an approach could not be implemented here, as it is computationally intensive, and it would be impossible to analyze the large number of simulated and experimental data sets generated here.

The results of spatial subsampling generated using the experimental data typically mirrored those of the simulated data, although a higher fraction of experimental time series exhibited early warning signals compared to the simulated data (fig. 2a, 2b). Possibly, this derives from the two different critical transitions (a fold bifurcation in the simulated data and a transcritical bifurcation in the experimental data). Alternatively, it may result from the inherently more noisy and short length of experimental time series, and thus the trends observed in the simulated data, where low search efforts produce a large variation in significant Kendall  $\tau$  values (fig. 2), are not seen in the experimental results.

Additionally, we show that temporal subsampling can influence the frequency with which early warning signals are detected (fig. 2c, 2d). This is unsurprising given that many leading indicators of population collapse are based on the temporal fluctuations of a population approaching

a transition (Dakos et al. 2012a). When the simulated population data was subsampled, all five leading indicators performed relatively well, even when sampling occurred on average only once every four time steps, so that data had to be interpolated (mean frequency of sampling  $\sim 0.25$ ; fig. 2c). This relatively infrequent sampling may give hope that currently available long-term data sets could prove useful in detecting early warning signals, even when sampling has been sporadic through time. We note, however, that even with very low temporal sampling (e.g., 0.1 samples per time step), the length of the time series generated by our simulation study would be  $\sim 80$  data points long, potentially longer than most data available to conservation practitioners. When the shorter time series from Drake and Griffin (2010a) were used, it became apparent that even with frequent sampling (approaching 1), type II errors were often present, likely a function of the noisy and (relatively) short nature of the data. This suggests that temporal subsampling of wild populations is likely in many cases to be more detrimental than spatial subsampling.

Here we have demonstrated how two common sampling issues associated with wild populations (search efforts and the temporal frequency of sampling) affect the prevalence and strength of early warning signals. However, there are numerous other factors that can influence estimates of population abundances through time. For example, imperfect detection of individuals can have important consequences for the identification of population dynamics and the estimation of species distributions (e.g., Lahoz-Monfort et al. 2014); misidentification when closely related species co-occur may exacerbate the problem (Scott and Hallam 2003), while the experience of the surveyors employed to make population counts can alter abundance estimates (Alldredge et al. 2007). At larger spatial scales, migration and immigration are known to significantly alter the persistence of populations (Schaub et al. 2010), and while the autocorrelation of movement between populations has been recognized as a potential early warning signal of collapse (Dakos et al. 2011), it remains unclear how these fundamental processes may affect the detectability of early warning signals in single population time series such as those presented here.

A related problem is the spatial distribution of individuals, assumed to be random in the work presented here (hence the use of a binomial distribution for the generation of sighting events)—clearly a simplification of the complexities of real-world scenarios. Where populations are highly aggregated, targeted sampling of small spatial areas could allow for a higher proportion of the total population to be sampled than expected if individuals were randomly distributed. This would then be equivalent to higher levels of spatial sampling in the simulations presented here. Search effort could thus also be considered as equivalent



to the approximate proportion of a population sampled and not simply the spatial area searched.

In conclusion, we show in this article that when long time series data are available, leading indicators of population collapse can be robust to observation error, including relatively infrequent surveys and poor search efforts. Of the five leading indicators we test here, autocorrelation at the first lag, first-order autoregressive coefficient, and density ratio were the most robust to varying search efforts and temporal sampling. However, when search efforts are low and sampling is infrequent, these statistical early warning signals are often unreliable. Thus, we additionally suggest that future analysis should seek to implicitly or explicitly include observational uncertainty or approach these issues in alternate ways (e.g., Boettiger and Hastings 2012).

### Acknowledgments

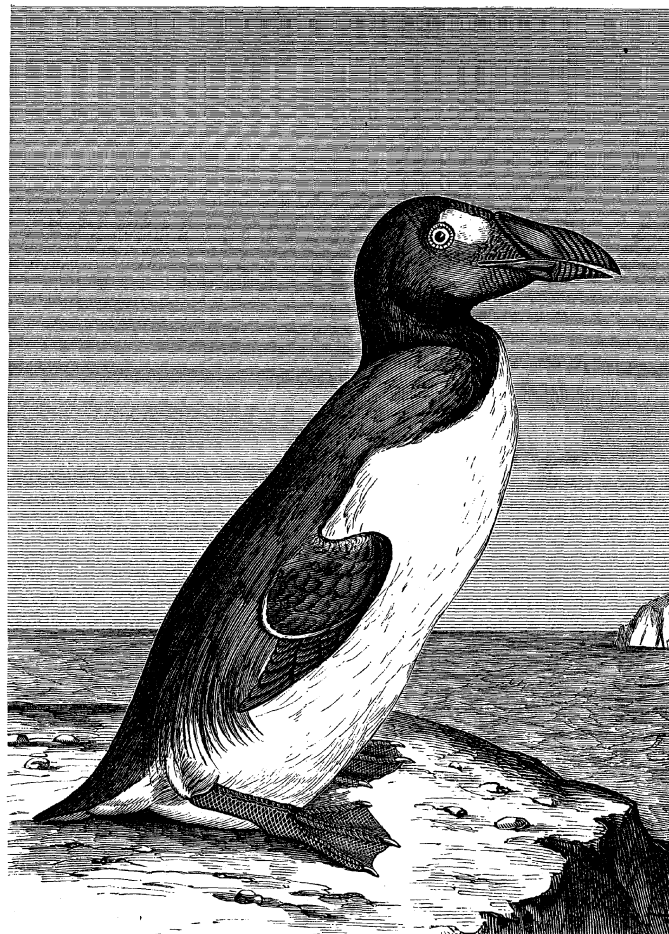
We thank T. Hefley and D. Kramer for their useful comments on earlier versions of this manuscript. This research was supported by the European Research Council Starting Grant (A.O.) and the Swiss National Science Foundation research grant (A.O.).

### Literature Cited

- Allredge, M. W., T. R. Simons, and K. H. Pollock. 2007. Factors affecting aural detections of songbirds. *Ecological Applications* 17: 948–955.
- Birkhead, T. 2014. Stormy outlook for long-term ecology studies. *Nature* 514:405.
- Boettiger, C., and A. Hastings. 2012. Quantifying limits to detection of early warning for critical transitions. *Journal of the Royal Society Interface* 9:2527–2539.
- Brock, W. A., and S. R. Carpenter. 2010. Interacting regime shifts in ecosystems: implication for early warnings. *Ecological Monographs* 80:353–367.
- Carpenter, S. R., J. J. Cole, M. L. Pace, R. Batt, W. A. Brock, T. Cline, J. Coloso, et al. 2011. Early warnings of regime shifts: a whole-ecosystem experiment. *Science* 332:1079–1082.
- Clark, J. S., and O. N. Bjørnstad. 2004. Population time series: process variability, observation errors, missing values, lags, and hidden states. *Ecology* 85:3140–3150.
- Clements, C. F., B. Collen, T. M. Blackburn, and O. L. Petchey. 2014. Effects of recent environmental change on accuracy of inferences of extinction status. *Conservation Biology* 28:971–981.
- Clements, C. F., N. Worsfold, P. Warren, B. Collen, T. Blackburn, N. Clark, and O. L. Petchey. 2013. Experimentally testing an extinction estimator: Solow's optimal linear estimation model. *Journal of Animal Ecology* 82:345–354.
- Collen, B., J. Loh, S. Whitmee, L. McRae, R. Amin, and J. E. M. Bailie. 2009. Monitoring change in vertebrate abundance: the living planet index. *Conservation Biology* 23:317–327.
- Contamin, R., and A. M. Ellison. 2009. Indicators of regime shifts in ecological systems: what do we need to know and when do we need to know it? *Ecological Applications* 19:799–816.
- Dai, L., D. Vorselen, K. S. Korolev, and J. Gore. 2012. Generic indicators for loss of resilience before a tipping point leading to population collapse. *Science* 336:1175–1177.
- Dakos, V., S. R. Carpenter, W. A. Brock, A. M. Ellison, V. Guttal, A. R. Ives, S. Kéfi, et al. 2012a. Methods for detecting early warnings of critical transitions in time series illustrated using simulated ecological data. *PLoS ONE* 7:e41010.
- Dakos, V., S. Kéfi, M. Rietkerk, E. H. van Nes, and M. Scheffer. 2011. Slowing down in spatially patterned ecosystems at the brink of collapse. *American Naturalist* 177:E153–E166.
- Dakos, V., M. Scheffer, E. H. van Nes, V. Brovkin, V. Petoukhov, and H. Held. 2008. Slowing down as an early warning signal for abrupt climate change. *Proceedings of the National Academy of Sciences of the USA* 105:14308–14312.
- Dakos, V., E. H. van Nes, P. D'Odorico, and M. Scheffer. 2012b. Robustness of variance and autocorrelation as indicators of critical slowing down. *Ecology* 93:264–271.
- Drake, J., and B. Griffen. 2010a. Early warning signals of extinction in deteriorating environments. *Nature* 467:456–459.
- Drake, J., and B. Griffen. 2010b. Data from: Early warning signals of extinction in deteriorating environments. *Nature* 467:456–459, Dryad Digital Repository, <http://dx.doi.org/10.5061/dryad.q3p64>.
- Ebert, D. 2005. Ecology, epidemiology, and evolution of parasitism in *Daphnia*. Introduction in *Daphnia biology*. National Center for Biotechnology Information, Bethesda, MD.
- Ford, J. H., M. V. Bravington, and J. Robbins. 2012. Incorporating individual variability into mark-recapture models. *Methods in Ecology and Evolution* 3:1047–1054.
- Franzetti, B., F. Ronchi, F. Marini, M. Scacco, R. Calmanti, A. Calabrese, A. Paola, et al. 2012. Nocturnal line transect sampling of wild boar (*Sus scrofa*) in a Mediterranean forest: long-term comparison with capture-mark-resight population estimates. *European Journal of Wildlife Research* 58:385–402.
- Green, R. H., and R. Young. 1993. Sampling to detect rare species. *Ecological Applications* 3:351–356.
- Hefley, T. J., A. J. Tyre, and E. E. Blankenship. 2013. Statistical indicators and state-space population models predict extinction in a population of bobwhite quail. *Theoretical Ecology* 6:319–331.
- Ives, A. R., and V. Dakos. 2012. Detecting dynamical changes in nonlinear time series using locally linear state-space models. *Ecosphere* 3:art58.
- Krkošek, M., and J. M. Drake. 2014. On signals of phase transitions in salmon population dynamics. *Proceedings of the Royal Society B: Biological Sciences* 281:20133221.
- Lahoz-Monfort, J. J., G. Guillera-Arroita, and B. A. Wintle. 2014. Imperfect detection impacts the performance of species distribution models. *Global Ecology and Biogeography* 23:504–515.
- Loh, J., R. E. Green, T. Ricketts, J. Lamoreux, M. Jenkins, V. Kapos, and J. Randers. 2005. The Living Planet Index: using species population time series to track trends in biodiversity. *Philosophical Transactions of the Royal Society B: Biological Sciences* 360:289–295.
- Muhlfeld, C. C., M. L. Taper, D. F. Staples, and B. B. Shepard. 2006. Observer error structure in bull trout redd counts in Montana streams: implications for inference on true redd numbers. *Transactions of the American Fisheries Society* 135:643–654.
- NERC Centre for Population Biology. 2010. The Global Population Dynamics Database. Ver. 2. <http://www3.imperial.ac.uk/cpb/data/bases/gpdd>.

- R Development Core Team. 2014. R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna. <http://www.r-project.org>.
- Schaub, M., A. Aebischer, O. Gimenez, S. Berger, and R. Arlettaz. 2010. Massive immigration balances high anthropogenic mortality in a stable eagle owl population: lessons for conservation. *Biological Conservation* 143:1911–1918.
- Scheffer, M., J. Bascompte, W. A. Brock, V. Brovkin, S. R. Carpenter, V. Dakos, H. Held, E. H. van Nes, et al. 2009. Early-warning signals for critical transitions. *Nature* 461:53–59.
- Scott, W. A., and C. J. Hallam. 2003. Assessing species misidentification rates through quality assurance of vegetation monitoring. *Plant Ecology* 165:101–115.
- ter Braak, C. J. F., A. J. Van Strien, R. Meijer, and T. J. Verstrael. 1994. Analysis of monitoring data with many missing values: which method? Pages 663–673 in E. J. M. Hagemeyer and T. J. Verstrael, eds. *Bird numbers 1992. Distribution, monitoring and ecological aspects*. Proceedings of the 12th International Conference of IBCC and EOAC, Noordwijkerhout, Netherlands. Statistics Netherlands, Voorburg.
- Wright, J. A., R. J. Barker, M. R. Schofield, A. C. Frantz, A. E. Byrom, and D. M. Gleeson. 2009. Incorporating genotype uncertainty into mark-recapture-type models for estimating abundance using DNA samples. *Biometrics* 65:833–840.

Associate Editor: Egbert H. van Nes  
Editor: Susan Kalisz



“The Great Auk [...] fed on fishes and marine plants, and laid either in the clefts of the rocks or in deep burrows a solitary egg [...]. The only noise it was known to utter was a gurgling sound. Once very abundant on both shores of the North Atlantic, it is now believed to be entirely extinct, none having been seen or heard of alive since 1844, when two were taken near Iceland.” From “The Great Auk” by James Orton (*The American Naturalist*, 1869, 3:539–542).